Oleg Gleizer
prof1140g@math.ucla.edu

Brendan Connelly
brendanconnelly@ucla.edu

Geometry - Introduction to Linear Algebra Packet 3

Preliminary - Warm Up

Problem 1 (Catalan warm-up: Dyck paths and parentheses)**.** Fix a positive integer n. A *Dyck path of semilength* n is a lattice path in the plane that

- starts at $(0, 0)$ and ends at $(2n, 0)$,
- uses steps $\vec{U} = (1, 1)$ (up-step) and $\vec{D} = (1, -1)$ (down-step),
- never goes below the x-axis (i.e. its y-coordinate is never negative).

A *balanced parentheses string of length* $2n$ is a string of n left parentheses "(" and n right parentheses ")" such that when you read from left to right, you never have more right parentheses than left parentheses.

1. For $n = 3$, list all balanced parentheses strings of length 6.
2. Describe a rule that turns a Dyck path of semilength n into a balanced parentheses string of length $2n$.
3. Describe the inverse rule that turns a balanced parentheses string into a Dyck path.
4. Prove that your two rules are inverses. You can then conclude there is a bijection between Dyck paths of semilength n and balanced parentheses strings of length $2n$.

. . . .

1. Matrices: what they are and how to compute with them

Note 1. A matrix is a rectangular array of numbers. You have probably seen matrices used to organize data or solve systems of equations. In this packet we will treat matrices as a *computational tool*: we will learn how to add, subtract, and multiply them.

Later (in Section 6–7) we will connect matrices to *linear transformations*. For now, think of a matrix as a machine that *takes an input list of numbers and produces an output list of numbers.*

Definition 1. *An* $m \times n$ matrix *is a rectangular array of real numbers with* m *rows and* n *columns:*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

We call a_{ij} *the* (i, j)-entry *of* A*.*

Note 2. Two matrices can be added only if they have the *same size.* Multiplication is different: it depends on *matching dimensions.*

. . . .

1.1. Adding and subtracting matrices

Definition 2. *If A and B are both $m \times n$ matrices, their sum is defined entry-by-entry:*

$$(A+B)_{ij} = A_{ij} + B_{ij}.$$

Similarly, $(A-B)_{ij} = A_{ij} - B_{ij}$.

Problem 2 (Compute $A+B$ and $A-B$)**.** Let

$$A = \begin{bmatrix} 2 & -1 \\ 4 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 3 & 5 \\ -2 & 7 \end{bmatrix}.$$

1. Compute $A+B$.
2. Compute $A-B$.

Definition 3 (Scalar multiplication)**.** *If A is an $m \times n$ matrix and $c \in \mathbb{R}$ is a scalar, the scalar multiple cA is defined entry-by-entry by*

$$(cA)_{ij} = c\,A_{ij}.$$

Problem 3 (Scalar multiplication)**.** Let

$$A = \begin{bmatrix} 1 & -3 & 2 \\ 0 & 4 & -1 \end{bmatrix}.$$

Compute $2A$ and $(-1)A$.

. . . .

1.2. Matrix multiplication

Note 3. Matrix multiplication is <u>not</u> entry-by-entry. Instead, it is designed so that multiplying by a matrix combines information from a whole row and a whole column.

Definition 4 (Matrix multiplication)**.** *Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. Then AB is the $m \times p$ matrix defined by*

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}.$$

In words: the (i, j)-entry of AB is the dot product of row i of A with column j of B.

Note 4 (Dimension rule)**.** The product AB is defined exactly when the number of columns of A equals the number of rows of B. If A is $m \times n$ and B is $n \times p$, then AB is $m \times p$.

Problem 4 (Which products make sense?)**.** For each pair, say whether the product is defined. If it is, state the size of the product.

1. $(2 \times 3) \cdot (3 \times 4)$
2. $(3 \times 2) \cdot (3 \times 2)$
3. $(4 \times 1) \cdot (1 \times 4)$
4. $(1 \times 4) \cdot (4 \times 1)$

Problem 5 (Compute a product)**.** Let

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}.$$

Compute AB and BA.

Problem 6 (Matrix times a vector)**.** Let

$$A = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}, \qquad \vec{v} = \begin{bmatrix} 4 \\ -2 \end{bmatrix}.$$

Compute $A\vec{v}$. *Note that a vector is a* $1 \times n$ *matrix*

. . . .

1.3. Key properties of matrix multiplication

Note 5. Matrix multiplication behaves in many ways like ordinary multiplication, but with one big warning:

In general, $AB \neq BA$.

Theorem 1 (Basic matrix laws)**.** Let A, B, C be matrices of compatible sizes, and let $r \in \mathbb{R}$.

1. (Associativity) $(AB)C = A(BC)$.

2. (Distributivity) $A(B+C) = AB + AC$ and $(A+B)C = AC + BC$.
3. (Scalar factors) $(rA)B = r(AB) = A(rB)$.

Note 6. We will not prove these laws in full generality here (the proofs are careful book-keeping with indices). Instead, we will practice using them correctly, and we will prove one important special fact.

Problem 7 (Practice: simplify using laws)**.** Assume all products make sense. Rewrite each expression without parentheses using the laws above.

1. $(A+B)C$
2. $A(B+C) - AB$
3. $(2A)(3B)$

Problem 8 (Non-commutativity: a concrete example)**.** Find two 2×2 matrices A, B such that $AB \neq BA$. (You may use the matrices from an earlier problem, or choose your own.)

Problem 9 (Identity matrix)**.** The $n \times n$ *identity matrix* is

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

1. Compute $I_2 \begin{bmatrix} a \\ b \end{bmatrix}$ and explain what it does.
2. Compute AI_2 and I_2A for a general 2×2 matrix $A = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$.
3. Conclude that I_2 behaves like a multiplicative identity for 2×2 matrices.

. . . .

1.4. One real-world example: tracking changes over time

Note 7. Here is a simple but powerful idea: matrices can keep track of updates. If a vector records a state (like amounts in categories), then a matrix can record how the state changes from one step to the next.

When you apply two updates in a row, you multiply the matrices. This is why matrix multiplication is useful.

Example 1 (Population / flow model (two categories))**.** Suppose a population is split between two regions, North and South. Let

$$\vec{v} = \begin{bmatrix} N \\ S \end{bmatrix}$$

record the number of people in each region.

Each month:

- 90% of North stays in North, and 10% moves to South,
- 20% of South moves to North, and 80% stays in South.

This update is described by the matrix

$$M = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}.$$

Indeed,

$$M\vec{v} = \begin{bmatrix} 0.9N + 0.2S \\ 0.1N + 0.8S \end{bmatrix}$$

is exactly the new population vector after one month.

After two months, the update is $M(M\vec{v}) = M^2\vec{v}$. After k months, it is $M^k\vec{v}$.

Problem 10 (Do two updates)**.** Let

$$M = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}, \qquad \vec{v}_0 = \begin{bmatrix} 1000 \\ 500 \end{bmatrix}.$$

1. Compute $\vec{v}_1 = M\vec{v}_0$. (Interpret your answer.)
2. Compute $\vec{v}_2 = M\vec{v}_1$.
3. Compute M^2. Check that $\vec{v}_2 = M^2\vec{v}_0$.
4. In one sentence: why does matrix multiplication naturally appear here?

2. Linear transformations: rules that respect addition and scaling

Note 8. Many real-world processes take an input vector and produce an output vector:

- physics: force $\mapsto$ acceleration, displacement $\mapsto$ new displacement
- geometry: rotate/reflect/stretch shapes
- economics: input bundle $\mapsto$ output bundle (in simple linear models)

We care about processes that behave nicely with adding and scaling, called linear transformations.

Definition 5. *Let V, W be vector spaces. A function $T : V \to W$ is a* linear transformation *if for all $\vec{u}, \vec{v} \in V$ and all $a \in \mathbb{R}$,*

$$T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v}), \qquad T(a\vec{v}) = aT(\vec{v}).$$

Example 2 (Geometry: stretch map in $\mathbb{R}^2$)**.** Define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(x, y) = (2x, y)$. This doubles horizontal distances but leaves vertical distances unchanged. One can check linearity directly:

$$T((x_1, y_1) + (x_2, y_2)) = T(x_1 + x_2, y_1 + y_2) = (2x_1 + 2x_2, y_1 + y_2)$$

and

$$T(x_1, y_1) + T(x_2, y_2) = (2x_1, y_1) + (2x_2, y_2) = (2x_1 + 2x_2, y_1 + y_2).$$

Similarly $T(a(x, y)) = aT(x, y)$.

Problem 11 (Check linearity (or find the failure))**.** Decide whether each map $T : \mathbb{R}^2 \to \mathbb{R}^2$ is linear.

1. $T(x, y) = (x + 1, y)$.
2. $T(x, y) = (x, y^2)$.
3. $T(x, y) = (x - y,\ x + y)$.

For any non-linear example, say which axiom fails (preservation of addition and/or scalar multiplication).

Theorem 2 (A linear map is determined by what it does to a basis)**.** Let $T : V \to W$ be linear and let $B = (\vec{b}_1, \ldots, \vec{b}_n)$ be a basis of V. If you know $T(\vec{b}_1), \ldots, T(\vec{b}_n)$, then you know $T(\vec{v})$ for every $\vec{v} \in V$.

Problem 12 (Prove the theorem)**.** Let $\vec{v} \in V$. Since B is a basis, write $\vec{v} = a_1\vec{b}_1 + \cdots + a_n\vec{b}_n$.

1. Use linearity to show $T(\vec{v}) = a_1 T(\vec{b}_1) + \cdots + a_n T(\vec{b}_n)$.
2. Explain why this means that knowing T on a basis determines T on all of V.

. . . .

3. Matrices: a compact way to store a linear transformation

Note 9. In $\mathbb{R}^n$, a linear transformation is often represented by a matrix. The idea is extremely simple: **the columns are the images of the basis vectors**.

Note 10 (Standard basis)**.** The standard basis of $\mathbb{R}^n$ is $\vec{e}_1 = (1, 0, \ldots, 0)$, $\ldots$, $\vec{e}_n = (0, \ldots, 0, 1)$.

Definition 6. *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be linear. The* matrix of T in the standard bases *is the $m \times n$ matrix*

$$A = \begin{bmatrix} T(\vec{e}_1) & \cdots & T(\vec{e}_n) \end{bmatrix}.$$

Then for every $\vec{v} \in \mathbb{R}^n$,

$$T(\vec{v}) = A\vec{v}.$$

Example 3 (A "graphics engine" map)**.** Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ satisfies

$$T(\vec{e}_1) = (2, 0), \qquad T(\vec{e}_2) = (1, 1).$$

Then

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}.$$

If $\vec{v} = (x, y) = x\vec{e}_1 + y\vec{e}_2$, then

$$T(\vec{v}) = xT(\vec{e}_1) + yT(\vec{e}_2) = x(2,0) + y(1,1) = (2x + y,\ y),$$

which matches $A\vec{v}$.

Problem 13 (From columns to formula)**.** Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be linear with

$$T(\vec{e}_1) = (-1, 2), \qquad T(\vec{e}_2) = (3, 0).$$

1. Write down the matrix A of T.
2. Compute $T(4, -1)$.
3. Describe in words what T does to the unit square with vertices $(0,0), (1,0), (0,1), (1,1)$.

Problem 14 (Drawing a transformation (geometry))**.** Let $T(x, y) = (x + y, y)$. (This is a *shear*.)

1. Compute $T(\vec{e}_1)$ and $T(\vec{e}_2)$.
2. Draw the image under T of the triangle with vertices $(0,0)$, $(2,0)$, $(0,1)$.
3. Is area preserved? (Guess from the picture; do not worry about proof yet.)

. . . .

4. A geometric application of coordinates

Problem 15 (Barycentric coordinates in a triangle (geometry))**.** Let $\vec{A}, \vec{B}, \vec{C}$ be three non-collinear points in $\mathbb{R}^2$. Any point $\vec{P}$ in the plane can be written as

$$\vec{P} = \vec{A} + \alpha(\vec{B} - \vec{A}) + \beta(\vec{C} - \vec{A})$$

for some real numbers α, β. (This is a coordinate system based at $\vec{A}$ with directions $\vec{B} - \vec{A}$ and $\vec{C} - \vec{A}$.)

You may remember Barycentric coordinates from last year

1. Explain why this is a span/basis statement in $\mathbb{R}^2$.

2. If $\vec{A} = (0,0)$, $\vec{B} = (2,0)$, $\vec{C} = (0,3)$, write $(1,1)$ in the form $\vec{A} + \alpha(\vec{B} - \vec{A}) + \beta(\vec{C} - \vec{A})$.

3. Describe the set of points $\vec{P}$ for which $\alpha \geq 0$, $\beta \geq 0$, and $\alpha + \beta \leq 1$.

4. Define $S : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$S(\alpha, \beta) \;=\; \alpha(\vec{B} - \vec{A}) + \beta(\vec{C} - \vec{A}).$$

Prove that S is a linear transformation.

5. Let $D = \{(\alpha, \beta) \in \mathbb{R}^2 \mid \alpha \geq 0,\ \beta \geq 0,\ \alpha + \beta \leq 1\}$. Show that the three vertices of D are $(0,0)$, $(1,0)$, and $(0,1)$, and compute

$$\vec{A} + S(0,0), \qquad \vec{A} + S(1,0), \qquad \vec{A} + S(0,1).$$

6. Conclude that the map $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(\alpha, \beta) = \vec{A} + S(\alpha, \beta)$ sends the triangle D exactly onto the triangle with vertices $\vec{A}, \vec{B}, \vec{C}$. (In other words, $\triangle ABC$ is the image of a standard triangle under an "almost linear" map.)